ORIGINAL PAPER

R-JaunLab: Automatic Multi-Class Recognition of Jaundice on Photos of Subjects with Region Annotation Networks

Zheng Wang[1,3] · Ying Xiao[4] · Futian Weng[1] · Xiaojun Li[4] · Danhua Zhu[6] · Fanggen Lu[5] · Xiaowei Liu[4] · Muzhou Hou[1] · Yu Meng[2]

Abstract

Jaundice occurs as a symptom of various diseases, such as hepatitis, the liver cancer, gallbladder or pancreas. Therefore, clinical measurement with special equipment is a common method that is used to identify the total serum bilirubin level in patients. Fully automated multi-class recognition of jaundice combines two key issues: (1) the critical difficulties in multi-class recognition of jaundice approaches contrasting with the binary class and (2) the subtle difficulties in multi-class recognition of jaundice represent extensive individuals variability of high-resolution photos of subjects, huge coherency between healthy controls and occult jaundice, as well as broadly inhomogeneous color distribution. We introduce a novel approach for multi-class recognition of jaundice to detect occult jaundice, obvious jaundice and healthy controls. First, region annotation network is developed and trained to propose eye candidates. Subsequently, an efficient jaundice recognizer is proposed to learn similarities, context, localization features and globalization characteristics on photos of subjects. Finally, both networks are unified by using shared convolutional layer. Evaluation of the structured model in a comparative study resulted in a significant performance boost (categorical accuracy for mean 91.38%) over the independent human observer. Our work was exceeded against the state-of-the-art convolutional neural network (96.85% and 90.06% for training and validation subset, respectively) and showed a remarkable categorical result for mean 95.33% on testing subset. The proposed network makes a performance better than physicians. This work demonstrates the strength of our proposal to help bringing an efficient tool for multi-class recognition of jaundice into clinical practice.

Zheng Wang and Ying Xiao contributed equally to this work.

✉ Muzhou Hou
  houmuzhou@sina.com

✉ Yu Meng
  mengyu1981@163.com

1   School of Mathematics and Statistics, Central South University, Changsha, Hunan 410083, China

2   Department of Gastroenterology and Hepatology, Shenzhen University General Hospital, Shenzhen, 518055, China

3   Science and Engineering School, Hunan First Normal University, Changsha, 410205, China

4   Gastroenterology Department of Xiangya Hospital, Central South University, Changsha, 410083, China

5   The Second Xiangya Hospital, Central South University, 410083, Changsha, China

6   Department of Gastroenterology, Hunan Provincial People's Hospital, Changsha, 410002, China

Introduction

Jaundice also known as an abnormal condition is a common finding for patients among at-risk populations (e.g., infectious mononucleosis [1], malaria [2], hepatitis [3], cirrhosis of the liver [4], gallbladder disease [5] and pancreatic cancer [6]). Jaundice represents a stained yellow color in the skin or sclera and is a byproduct of old red blood cells due to bilirubin metabolic disorders. Levels of total serum bilirubin (*TBil*) in blood are normally below 17.1 μmol/L, and levels over 17.1 μmol/L typically result in jaundice [7, 8]. Jaundice is further classified as two types: occult jaundice (clinical) and obvious jaundice (subclinical). Obvious jaundice can be confirmed by finding levels of *TBil* over 34.2 μmol/L that can cause a yellowish pigmentation, in particular, the skin and whites of the eyes. Level between 17.1 μmol/L and 34.2 μmol/L causes occult

jaundice, which would not lead to yellowish staining or mucous membranes by bilirubin [9]. Noticeably, causes of jaundice vary from non-serious to potentially fatal [10]. In this work, a multi-class recognition of jaundice was developed to provide an optimal diagnostic schedule, which can control the heme load imparted to the circulation early and make substantial diagnosis as well as prognosis result of underlying cause [11].

Existing works in jaundice recognition are driven by the estimation of bilirubin levels on conventional chemical and biological procedures, such as urine test, serum evaluation and liver function check which are time-consuming, invasive and expensive. In the last decade, several automatic and semi-automatic methods have been developed to classify two general categories (i.e. healthy controls and obvious jaundice) [12–16]. Some methods use non-invasive jaundice measurement [12, 17–20] or sequential Bayesian model [21] or with computer version [22]. Manual multi-class recognition for jaundice (including occult jaundice) on photos of subjects is a challenging task. There are three crucial challenges: (1) skilled individuals have rich experience, which are difficult to inherit; (2) the task is time-consuming, invasive and expensive that makes it infeasible for large-scale datasets in primary-level hospitals and clinical practice; and (3) over-fatigue might bring about missed diagnosis that increases harmful level of disease which may result in casualty. Three classes of photos of subjects of jaundice are presented in Fig. 1. Consequently, we proposed a novel model for multi-class recognition of jaundice, which can alleviate the heavy workloads of skilled individuals and provide them with more human service.

Recently, artificial intelligence [23–26] has remarkable improvement of image classification [27–30], object detection [31–37] and recognition techniques [38–40]. [38] introduced various techniques used for object recognition system. [41] proposed a framework to uncover knowledge in a database, bringing light to disguise patterns which can help in credible decision making. [33] presented an approach that can efficiently detect objects in an image while simultaneously generating a high-quality segmentation mask for each instance. Current approaches train models in multi-stage training algorithm (pipeline method) that provides a jointly learning algorithm to classify object proposals. Multi-class recognition of jaundice requires the accurate prediction of jaundices which is a more challenging task due to the complexity, which creates two serious obstacles. (1) eye object must be localized on photos of subjects (termed as 'eye with annotation') and (2) these candidates must be precise classified, i.e., precise multi-class recognition of jaundice. In this paper, we study a single training model that uses joint learning algorithm to localize eye object with possible annotations for training to provide an accurate and reliable solution and at the same time compromises speed, accuracy or simplicity. We train, validate and test our work on photos of subjects that could be useful in optimal values learning of millions weights in our deep networks. We test and process comparison with different strategies and network architectures. Results of experiment and comparisons provide not only its performance outperforming an independent human expert, but also suggest our best performing network significantly differing from the state-of-the-art classifier with RAN.
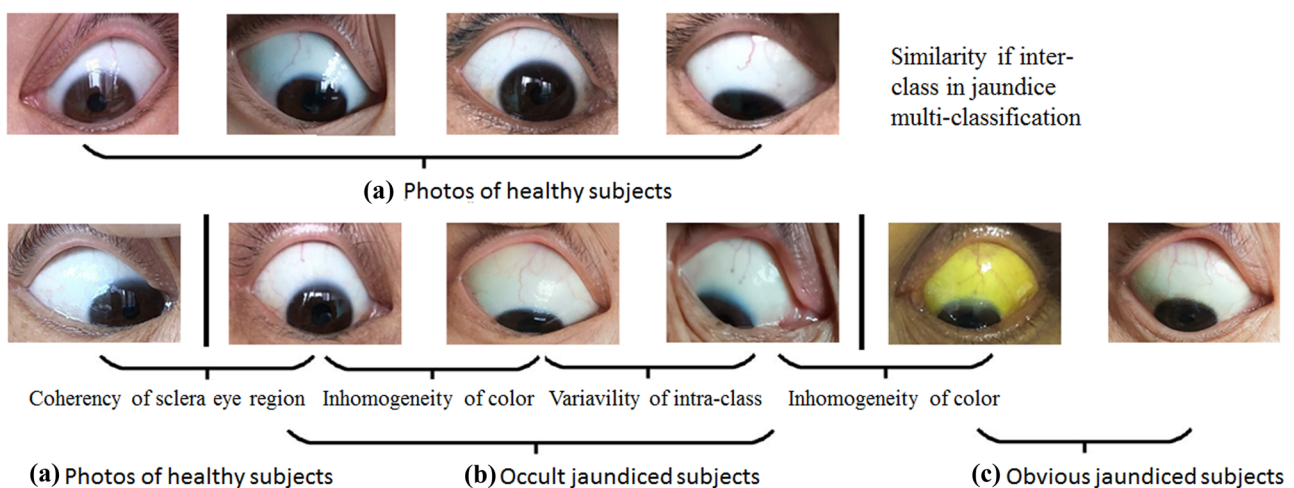

Fig. 1 Three classes of jaundice photos of subjects. There is great insight into photos of subjects and appear the broad variability, high coherency of sclera region, as well as extensive inhomogeneity on photos of subjects. These photos of subjects were all collected at original photo from smartphone under an image data acquisition and retrospective study protocol

Methods

Methods that were used in this study were applied to 134 healthy and 268 jaundiced subjects sclera images. The study was approved by the Ethics Committee of the 2nd Xiangya Hospital, Central South University. Informed consent was obtained from the subjects for the publication of the article. These data were collected from the 2nd Xiangya Hospital and Hunan Provincial People's Hospital. Outpatients and inpatients in the Department of Gastroenterology and Hepatology who had just received serum bilirubin assay were selected as subjects according to their test results. Those with serum bilirubin level of 1.7-17.1μ mol/L and more than 17.1μmol/L were classified as normal control group and jaundice group, respectively. Class between the two groups was defined as inter-class jaundice. Subjects with serum bilirubin level of 17.1-34.2μ mol/L and more than 34.2 μmol/L were classified as occult jaundice group and obvious jaundice group, respectively. Class between the two groups was defined as intro-class jaundice. During the data collection phase, subjects were asked to sit upright, upper sclera was exposed, and the camera was about 25cm away from the face. Using 12 MP

iPhone 8 Plus smartphone and no flash when using this smartphone camera, subjects′ upper sclera photos were taken and results were obtained in a specific room with a good lighting condition. The resolution value of the obtained images was 96 dpi, and for the other stage, these images were transferred to Dell xps8930 server and image processing methods and machine learning were used on these images, respectively.

The proposed model recognizes multi-class jaundice and is based on learning, data-driven approach. The R-JaunLab utilized learning approach, which can automatically annotate eye region and learn hierarchical feature representation. The R-JaunLab performed a data-driven approach by using the data augmentation manner, which reinforces the multi-class manner to yield efficient performance and more reliability. Hence, the overall approach proposes an end-to-end model to recognize varied jaundice. Fig. 2 illustrates the R-JaunLab architecture that reaches a breakthrough for the above-mentioned obstacles by employing region annotation network and hierarchical feature representation. The main contributions of R-JaunLab are summarized in the following insights:
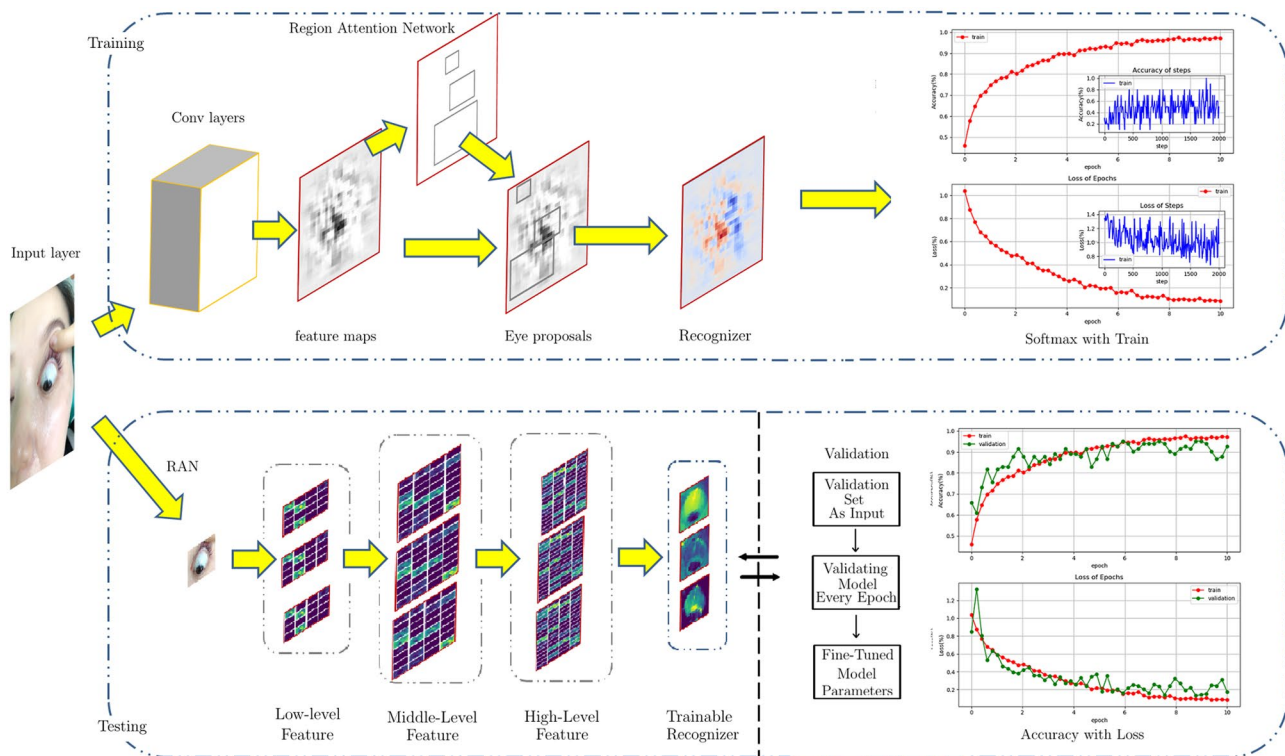

Fig. 2 Illustration of the proposed workflow. The proposed approach consists of three phases, such as training phase, validation phase and testing phase. The training stage learns the sufficient feature representation, which localizes the eye region and recognizes the proposals to classify jaundice. The purpose of the validation helps to optimize and fine-tune parameters in each epoch. The testing phase assesses the achieved performance of the R-JaunLab

- R-JaunLab model is proposed for the multi-class recognition of jaundice, which is processed in end-to-end manner. The model achieves a significant accuracy and suggests that it is potential to greatly reduce the skilled individuals workload and assist an early therapeutic schedules. Automatic recognition of multi-class jaundice has more values for clinicians and provides an reliable solution than binary classification in jaundice diagnosis, therapy and prognosis. Nevertheless, this domain has not ever investigated in the literature.
- R-Jaundice is composed of two approaches. The first approach employs a region annotation network by using deep convolutional structure, and the second approach is the jaundice recognizer that classifies jaundice by using the proposed eye candidates. Using the possible 'annotations' training of neural networks, the RAN module tells the R-JaunLab which proposes to classify.
- An efficient jaundice recognizer is proposed to compute similarities, localizations, context and globalization of feature representation on photos of subjects and employs prior knowledge between intro-class and inter-class of jaundice. Therefore, the R-JaunLab has excellent capabilities of feature learning that can distinguish more features under photos of subjects.

In Section 2.1 we introduce the architectures and characteristics of RAN. In Section 2.2 we develop a novel model for multi-class recognition of jaundice with shared feature from RAN.

Region Annotation Network

A region annotation network (RAN) takes an photos of subjects (of any size) as input and output for an eye region proposal, each with an Bland-Altman score [42, 43]. We model this process with a encoding–decoding CNN [44] to generate eye region proposal, which shares with the multi-class recognition of jaundice network. Region annotation network is a fine pixel-wise detection from eye region annotations.

According to the structure, the RAN consists of two main parts: the contracting encoding and expansive decoding units. The nonlinearity applied to the basic convolutional operations is a rectified linear units (ReLU [45], it is computed as Eq. 1) to prevent the vanishing gradient problem in both parts of the network. In the encoding path, 2×2 max-pooling operations (introduced in [46]) are performed to downsample the image by using maximum activation function. In the decoding path, we use 2×2 upsampling layers. Skip connection combined downsampled features from encoding path with the corresponding upsampled output from decoding path.

$$ReLU(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases} \tag{1}$$

We apply a minimum bounding-box approach [32] for computation of eye proposals on the last convolutional layer, which are shared for both networks. The minimum bounding-box approach takes nonzero region, which maps a corresponding eye region. Each minimum bounding box segment has proposal of eye region for multi-class recognition of jaundice. For minimum bounding box algorithm (Eq. 2.1), we exploit the parameterizations of the four coordinates as follows [32]:

$$\begin{aligned} e_x &= (x - x_a)/w_a, & e_y &= (y - y_a)/h_a \\ e_w &= \log(w/w_a), & e_h &= \log(h/h_a) \\ e_x^* &= (x^* - x_a)/w_a, & e_y^* &= (y^* - y_a)/h_a \\ e_w^* &= \log(w^*/w_a), & e_h^* &= \log(h^*/h_a) \end{aligned} \tag{2}$$

where x and y are the eye region's center coordinates, and w and h are the eye region's width and height, respectively. x, x_a and $x*$ denote the predicted eye region, eye region and ground-truth label, respectively. Eq. 2.1 is minimum bounding box algorithm segment of proposal per eye region. For potential eye region on each sample of photos of subjects, Eq. 2.1 assigns a binary label, when the potential eye region has the probability higher than 0.7. It is remarkable that nonzero values of the binarized label are a positive annotation for eye region and a negative annotation relevant to a non-eye region, as the probability of Eq. 2.1 is lower than 0.3.

We minimize an objective function following the binary cross-entropy (BCE) loss [47] in R-JaunLab. Our loss function for a region attention proposal network is defined as Eq. 3:

$$L_{ran}(p_i) = -\sum_{i=i}^{n} \hat{p}_i \log p_i + (1 - \hat{p}_i) \log(1 - p_i) \tag{3}$$

This is a loss function between probabilities, where $\hat{p}$ represents the distribution of the ground-truth and p is the probable distribution of eye region proposal. The loss is zero only if p_i and $\hat{p}_i$ are equal; otherwise, loss is a positive number. Moreover, the smaller the probability difference, the smaller the loss. The region annotation network provides several advantages for detection tasks:

(1) This model performs the global location and context at the same time.
(2) It is suitable for very few training samples and provides remarkable performance for detection tasks.
(3) An end-to-end successive processes the entire photo in the forward propagation and directly produces shared computation of minimum bounding box.

Sharing Features for RAN and Jaundice Classification

Convolutional neural networks have shown substantially higher accuracy of image classification on the ImageNet competition of large-scale visual recognition challenge (ILSVRC) [48, 49]; there is an increasing interest in the field of medical imaging [50–56]. Nevertheless, slow exploitation of CNNs in biomedical community is partly because training and testing on biomedical image datasets are relatively insufficient labeled that rest upon large skilled individuals. In this paper, transfer learning [54, 55, 57, 58] is capable of breaking through the very little training data problem.

For the multi-class recognition of jaundice, we employ a bottleneck features transfer learning of the 16-layer model from VGG16 [29], which consists of sequential 13 convolutional layers and 3 fully connected layers. The last 3 fully connected layers in VGG16 were replaced with a global average pooling [59], a layer with 100 outputs and a multi-class softmax classifier [60]. (Modified VGG16 is termed *jaundice recognizer* and shown in Fig. 3.) The loss function of the model for the multi-class probability distribution is as follows in Eq. 4, where $\hat{c}$ is a true probability distribution, i.e. one-hot coding for labels, and c is the probability distribution of eye region proposal, i.e. the result of softmax classifier output.

$$L_{jc}(c_i) = -\sum_{i=1}^{n} \hat{c}_i \log c_i \qquad (4)$$

In the training stage, the weights of the 13 convolutional layers serve as feature extractor. Training of the *jaundice recognizer* employs the end-to-end strategy, which extracts automatically discriminative, semantic and hierarchical features from low level to high level. The *jaundice recognizer* takes into account the relationship between intro-class and inter-class of the predicted proposals that can overcome the barriers from various photos of subjects. Particularly, the similarities of photos are measured as by the distance of feature space.

Both RAN and *jaundice recognizer* are trained independently. Therefore, R-JaunLab develops a algorithm that describes a unified network composed of RAN and *jaundice recognizer* with shared convolutional layer (as shown in training block in Fig. 2), rather than learning separate networks. In Fig. 2, the RAN and *jaundice recognizer* models are fused into one networks during training. In each iteration of RAN network, the forward propagation generates eye region proposals, which are also the input data of *jaundice recognizer*. The shared convolutional layer in *jaundice recognizer* accepts the predicted eye region as input and convolutes. Therefore, both models are combined by the shared convolutional layer and form a unified network, i.e. R-JaunLab.
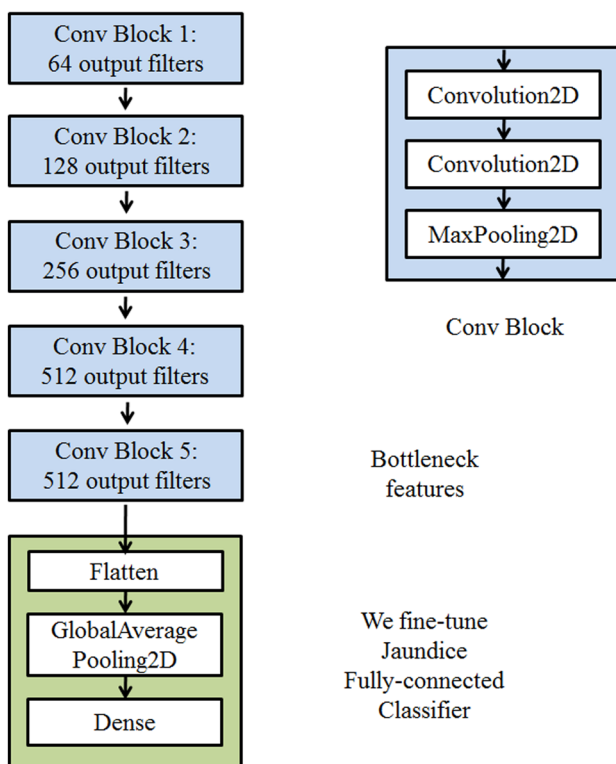
Workflow Overview

General workflow of R-JaunLab contains three top-down steps, as shown in Fig. 2. The detailed stages are described as follows:

– **Training stage:** The training stage learns to extract proposals of eye region by using region attention network on photos of subjects and classifying feature characteristics of different classes. After importing three classes patient-wise images, the R-JaunLab first learns to predict minimum bounding boxes by RAN and share the relevant eye region proposals that propagates as input for *jaundice recognizer* network. This recognition network is a pre-trained VGG16 model, which is initialized with an ImageNet, and fine-tuned with end-to-end strategy for the multi-class recognition task. During the training, the recognition network learns the localization, context, globalization and hierarchical features and optimize the relevant parameters. The assembled features propagate into softmax classifier. The results of the three classes are transmitted to constrained loss function, which maximize the features of inter-class and minimize the features of intra-class.

Fig. 3 Layer structures of the modified VGG16 neural networks used in the study (called jaundice recognizer). Blue: weights were frozen and training is unavailable; Green: final classifier and is trainable

- **Validation stage:** The validation process fine-tunes hyper-parameters and avoids overfitting. The best model is preserved for testing. The validation stage optimizes the multi-class recognition model on the photos of subjects, as demonstrated in validation boxes of Fig. 2.

- **Testing stage:** The goal of the testing stage is to assess the performance of the R-JaunLab. R-JaunLab process evaluation of multi-class recognition of jaundice and is shown in the testing box of Fig. 2. The first step performs eye region proposals with shared convolutional layer and then feature hierarchy (e.g. simple features, obvious features and discriminative features) can be learned or extracted via repeated iterations and fed into a trainable recognizer.

Finally, we comprehensively evaluate two training strategies. The first one directly trains R-JaunLab on patient-sourced photo dataset, i.e. training the *'R-JaunLab from scratch'*. Another one employs transfer learning technique that initialized with pretrained ImageNet and then fine-tunes it on photos of subjects dataset, i.e. training the *'R-JaunLab from transfer learning'*. The *'R-JaunLab from scratch'* performed worse in both accuracy and cross-entropy loss. *'R-JaunLab from transfer learning'* is more valuable and is chosen as the final strategy. Moreover, the number of training iteration was 50 due to the optimal accuracy from the validation and test set.

Implementation Details

The CNN models are exploited to perform the training on a Dell xps8930 server, which contains hexacore 3.20 GHz processor, 16 GB RAM and one NVIDIA GeForce GTX 1070 video card. This work was implemented in Python by using the Keras framework, which back with a TensorFlow.

Data Augmentation: As for our tasks, there are relatively small data available; we utilize excessive technique for data augmentation [61, 62] through four types of transformation. The transformation operators are according to the matrix (Eq. 5) as follows:

$$
\begin{aligned}
T_{flip} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
T_{zoom} &= \begin{bmatrix} z_x & 0 & 0 \\ 0 & z_y & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
T_{shift} &= \begin{bmatrix} 1 & 0 & r \times s_x \\ 0 & 1 & c \times t_y \\ 0 & 0 & 1 \end{bmatrix} \\
T_{rotation} &= \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}
\tag{5}
$$

z_x, z_y of T_{zoom} was uniformly sampled from (0.3, 1). T_{shift} processes a shift transformation, where r and c are the row number and the column number of the photo, respectively. s_x and t_y from T_{shift} matrix are sampled uniform in the range (-0.2, 0.2). θ in $T_{rotation}$ is uniformly sampled from the range (-5, 5).

Learning Rate Policy: The RAN is an end-to-end model and trained by using the root mean square prop (RMSprop [63]) algorithm, which changes adaptive the learning rate to speed up the training procedure and is initialized with learning rate η of 0.0001. The mini-batch size of RAN is 32 per photos of subjects. We adopt a momentum term γ of 0.9. The form of RMSprop (Eq. 6) is derived as follows:

$$
\begin{aligned}
E[g^2]_t &= \gamma E[g^2]_{t-1} + (1-\gamma)g_t^2 \\
\theta_{t-1} &= \theta_t - \frac{\eta}{\sqrt{E[g^2]_t + \epsilon}}g^2
\end{aligned}
\tag{6}
$$

For training the *jaundice recognizer* and learning weights, we employ the stochastic gradient descent (SGD [64, 65]) algorithm, mini-batch size of 32 and a cross-entropy cost function. SGD is computed as Eq. 7:

$$
\begin{aligned}
J_{train}(\theta) &= \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2 \\
\theta_j &:= \theta_j - \eta \frac{1}{m} \frac{\partial}{\partial \theta_j} J_{train}(\theta) \\
&:= \theta_j - \eta \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})x_j^{(i)}
\end{aligned}
\tag{7}
$$

Here, η is the learn rate and sets as 0.0001 for mini-batches on photos of subjects. We employ 0.9 and 0.0005 for momentum and weight decay, respectively [23].

Loss Function: High accuracy of loss for multi-class recognition of jaundice is critical in this study. We define loss function in R-JaunLab for an photos of subjects as Eq. 8:

$$
L(p_i, c_i) = \frac{1}{N_{ran}} \sum_i L_{ran}(p_i, \hat{p}_i) + \lambda \frac{1}{N_{jc}} \sum_i p_i L_{jc}(c_i, \hat{c}_i)
\tag{8}
$$

Here, a mini-batch takes the i as index and p_i of the ith predicted probability of an eye region. $\hat{p}_i$ is the ground-truth label. When the prediction is positive, $\hat{p}_i$ is 1. If the prediction is negative, $\hat{p}_i$ is 0. c_i is a class and represents the jaundice type (such as, benign, occult jaundice or obvious jaundice). $\hat{c}_i$ is prediction associated with a positive eye region. The loss L_{ran} refers to Eq. 3 in subsection 2.1, and the loss L_{jc} refers to Eq. 4 in subsection 2.2. The term p_i means the multi-class recognition loss, which is activated only

for positive eye region ($p_i = 1$) and is disabled otherwise ($p_i = 0$). The outputs of the *ran* and *jc* layers include p_i and c_i respectively.

The two terms of Eq. 8 are normalized by N_{ran} and N_{jc}. λ is weight factor and controls the trade-off of *ran* and *jc* losses. The *ran* term is normalized with $N_{ran} = 512$ and the *jc* term is normalized with $N_{jc} = 112$. By default we control $0 \geq \lambda \leq 1$. By optimizing cross-validation, the weight term λ is finally optimized with 0.5. Equation 8 is optimized by a SGD with momentum 0.9.

Results

Characteristics of Photos of Subjects: Initially, we obtained frontal photos of subjects (one image per subject), which passed image quality review and were used for the proposed model. The images were collected from the Second Xiangya Hospital of Central South University and labeled according to bilirubin levels testing (i.e. indicator of *TBIL*), which is conventional chemical and biological procedures (Fig. 4 text side). The left side of Fig. 4 uses yellow boxes on photos of subjects to localize the region of interest, which is always solely the eye region.

The subjects include both genders and the age of patient vary from 14 to 82 years. The dataset provides 402 patient-wise images and consists of 134 healthy individual (H), 90 occult jaundice (OC) and 178 obvious jaundice (OB) through *TBIL* level. Images are of RGB, three-channel and 4600 × 3400 size. Table 1 shows the photos of subjects distributions of three classes for each jaundice diagnosis category.

Reliability and Generalization: To promote reliability of the results, the whole photos of subjects are split into three groups according to patient-wise: training subset, validation subset and testing subset (as listed in Table 1). The ratio of three groups is 7:2:1 on the photos of subjects according to patient-wise. During the training phase, the training subset is used for the R-JaunLab model and parameters of different neurons are optimized. The validation subset is used to test the generalization capabilities of our model. The testing subset is used to evaluate the multi-class recognition accuracy of jaundice and reliability for clinicians.

In this study, the photos of subjects are augmented by applying elastic deformation and resolve the very little training data problem. The data augmentation is done on the training phase due to the standard method in machine learning community [62].

Threefold of the photos of subjects are non-overlapping. The results of all experiment are computed accuracy for mean with an average standard deviation and are evaluated using auxiliary performance metrics, named precision, sensitivity as well as specificity [66–69]. To evaluate the generalization, R-JaunLab and RAN are compared with other state-of-the-art classifiers and are validated on experiments of the multi-class jaundice recognition.

Recognition rates: Evaluating the performance of the proposed model on photos of subjects, there are two computing methods for reporting the results [70]. An independent test of validation subset and testing subset was employed to compare the recognitions of R-JaunLab (i.e., *image-level*) with the recognitions made by experts (i.e., *expert-level*). Six human experts with significantly clinical experience in the Second Xiangya Hospital of Central South University were instructed to make the recognitions on the photos of subjects.

As a start, the recognition rate is evaluated at *expert level*. N_e is the number of photos of subjects. If N_{er} photos are correctly recognized, expert score can be defined as Eq. 9:

$$S_{expert-level} = \frac{N_{er}}{N_e} \qquad (9)$$

Second, we evaluate the recognition rate at the *image level* that provides an approach to estimate solely the multi-class recognition accuracy of the R-JaunLab. N_p is the number of photos of subjects from the validation subset or testing subset. Let N_{pr} be correctly recognized number on patient-sourced photos, then the recognition rate of *image level* is depicted as Eq. 10:

$$S_{image-level} = \frac{N_{pr}}{N_p} \qquad (10)$$

Fig. 4 Acquired photos of subjects with bilirubin level test information of the same day

Table 1 Patient-wise distribution of photos of subjects dataset

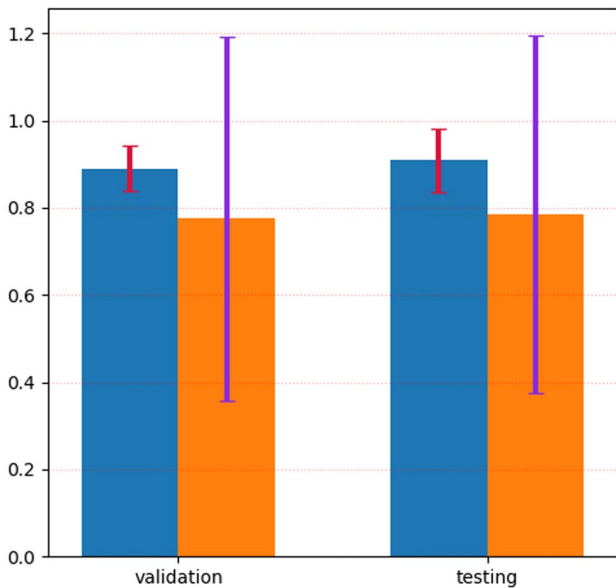| Photos of subjects	Jaundice class		
	Healthy	Occult	Obvious
Training Set	93	63	124
Validation Set	27	18	35
Testing Set	14	9	19
TBIL($\mu mol/L$)	≤ 17.1	$(17.1, 34.2)$	≥ 34.2

Fig. 5 Multi-class recognition of jaundice performance among *expert-level* and *image-level*. The experiments from both are processed on the same validation subset and testing subset. Accuracy for mean of *image level*: blue; accuracy for mean of *expert level*: orange; mean standard deviation of *image level*: red; mean standard deviation of *expert level*: violet

We compute the average score of the both recognition experiments (*expert-level* and *image-level*), which are assigned on validation subset and testing subset. The score of multi-class recognition of jaundice achieved a remarkable high performance and is reliable (as shown in Fig. 5). The total score for mean is 78.04% for the *expert level* and *image level* is 89.84%. The validation set (score for mean with 89.04% and 5.25% standard deviation for mean) and testing set (score for mean with 90.83% and 7.19% standard deviation for mean) of the proposed model have almost the same score. The proposed method demonstrates similar

recognition rates and strong generalization capabilities of novel network structure.

Performance: The performance of *'R-JaunLab from scratch'* and *'R-JaunLab from transfer learning'* is illustrated in Fig. 6 and demonstrates that transfer learning strategy is promising better than training from scratch. Training and validation from scratch roughly converge both to a low accuracy, which will be under-fitting. In the plot, the R-JaunLab training could benefit from transfer learning.

After 50 epochs (iterations through the whole dataset), the training and validation results of *'R-JaunLab from transfer learning'* show that it possesses not only excellent performance in both accuracy and cross-entropy loss but also uniform convergent, which demonstrates that the *'R-JaunLab from transfer learning'* has generalization and the capability to avoid over-fitting and under-fitting [72].

Data augmentation is used in R-JaunLab to enhance the small dataset and achieved significant performance as shown in Fig. 7. In comparison with the two datasets (Aug and Raw), the average accuracy of the Aug is 95.01% and 92.35%, while Raw is 76.5% and 75.5% (for training set and validation set, respectively), which demonstrates that augmentation available photos of subjects can meet the requirement of the model.

In order to differentiate the effect of input shape size between 112×112 and 224×224, we also performed a 'lower resolution input model' recognizing between the same three categories, i.e., we just change the input resolution without further adjustment to R-JaunLab. In each case, the experiments were trained until convergence, and then their performance was evaluated on the validation subset and testing subset of the photos of subjects. Resulting resolution for this procedure is demonstrated in Fig. 8. Using the same photos of subjects, the 'lower resolution input model' achieved an accuracy of 95.02% in training stage, 92.35% in validation stage and 95.33% in testing stage, respectively. Our 'lower resolution input

Fig. 6 The comparison between transfer learning and from scratch of R-JaunLab on training and validation subset. Accuracy and loss are plotted against the epoch during the length of training over the course of 10,000 steps. Training from transfer learning: red; validation from transfer learning: green; training from scratch: blue; validation from scratch: yellow
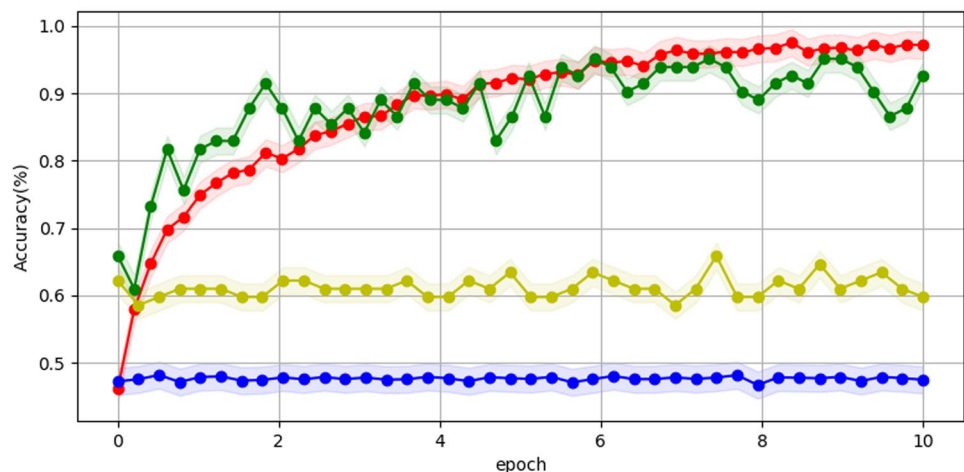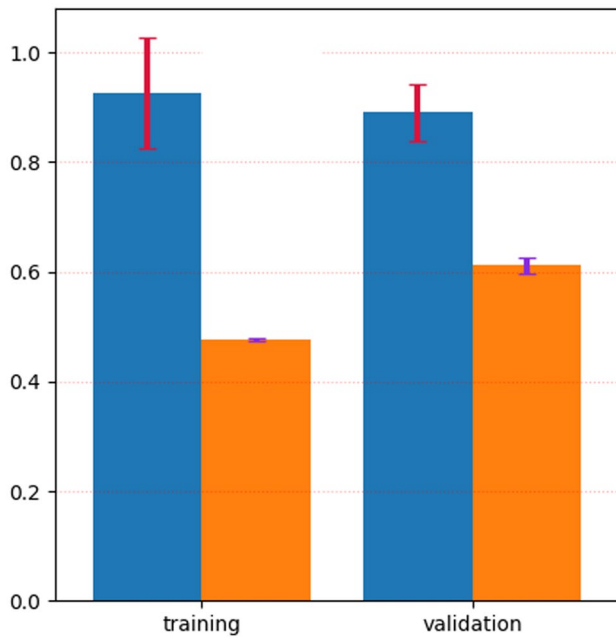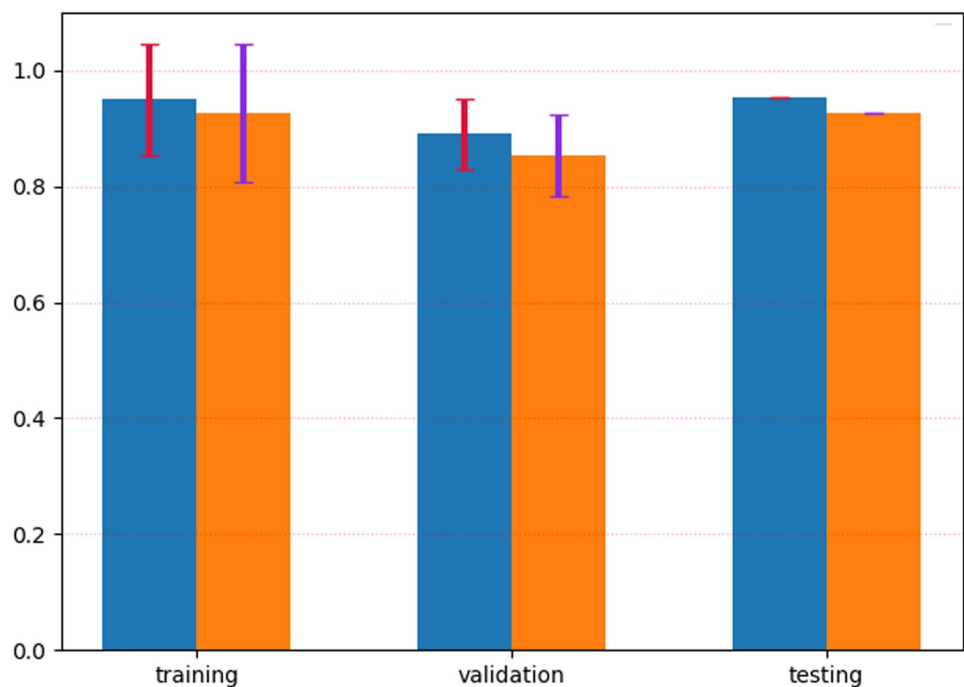
Fig. 7 R-JaunLab comparisons on photos of subjects between the raw photos (Raw) and augmented photos (Aug). Accuracy for mean of Aug: blue; accuracy for mean of Raw: orange; average standard deviation of Aug: red; average standard deviation of Raw: violet

model' improved about 2-3% accuracy, which shows that the capability of characteristic learning with 112×112 input resolution for R-JaunLab is better than input resolution with 224×224. Hence, 112×112 is an optimal strategy of input resolution and is used to reduce the computational costs in this study.

The training phase processed about one hour and two hours thirteen minutes under the whole photos of subjects and photos of subjects with augmentation data, respectively. Different shape sizes of input resolution took respectively about 50 minutes for 112×112 and two hours ten minutes for 224×224 on the training subset. Data-augmented operators were used in training and validation phase and were executed on Python 3.6.5. In the test stage, a single mini-batch took about 0.04s on the raw testing subset.

Discussion

It is the first time that occult jaundice is suggested for early diagnostic control, which inspires multi-class recognition of jaundice challenge. In this work, we proposed R-Jaundice for multi-class recognition of jaundice on photos of subjects. R-JaunLab learns higher-level discriminating features with RAN and achieves reliable and accurate recognition scores. By validation of the challenge of the small dataset, the performance in the above section determined that our model is capable to acquire the eye region and has the remarkable performance in multi-class recognition of jaundice. Although the photos of subjects have high resolution that bring about inter-class and intra-class challenges in the multi-class recognition task, the distinguished power of the R-JaunLab is better than the state-of-the-art CNNs. Furthermore, R-JaunLab has stable performance in multi-class recognition of jaundice on photos of subjects. The model demonstrates great value of applicability in clinical practice of jaundice. Since jaundice recognition faces a

Fig. 8 R-JaunLab comparisons between the different shape sizes of the input resolution (i.e., 112×112 and 224×224, respectively) The computational cost of both is almost constant. Input resolution with 112×112: blue; input resolution with 224×224: orange

time-consuming [73–75], invasive [76, 77] and expensive procedure [21, 78], our work could provide an automated multi-jaundice recognition system, which offers scientific, objective and concrete indexes.

Table 2 depicts several comparisons with popular classifiers and illustrates that the R-JaunLab outperforms other classifiers and marked in bold type. The Xception network proposed by [71] is a deep learning with depthwise separable convolutions and used for JFT (an internal Google dataset for large-scale image classification dataset) [79] with high accuracy. The Xception with RAN achieves about an accuracy with training set of 87.84% and with validation set of 83.29% in the multi-class recognition of jaundice photos of subjects. ResNet framework [27] is a residual learning CNNs proposed by Kaiming He and took the first place of localization and classification in the ILSVRC15 (ImageNet large-scale visual recognition challenge 2015). ResNet50 with RAN achieves an accuracy of 90.61% on training set and 82.13% on validation set. Inception model [30] proposed by Christian Szegedy is a convolutional network with inception architecture and reported 3.5% top-5 error and 17.3% top-1 error on ILSVRC12 classification challenge validation set. InceptionV3 with RAN achieves about an accuracy of 92.90% and 85.51% (for training set and validation set, respectively). VGG architecture [29] is a traditional CNN proposed by Karen Simonyan and secured the first and the second places in the ILSVRC14 [49] localization and classification tracks, respectively. VGG19 with RAN achieves an accuracy of 88.76% in training phase and 77.48% in validation phase. Table 2 demonstrates also results of other performance metrics (e.g., respective precision, sensitivity and specificity) on training set and validation set. The scores of the proposed work exceeded the state-of-the-art classifiers under auxiliary performance metrics.

In the testing phase, we performed an experiment to identify predicted result of our model. This test suggests that the performance of R-JaunLab outperforms the state-of-the-art prior models in multi-class recognition task of jaundice, as shown in Table 3. In the total accuracy for mean aspect, the proposed method is about 4% higher than the best performance of the existing methods and promoted to 95.33%. In particular, obvious jaundice was recognized correctly by experiment in 100% of all the photos of subjects, and healthy controls achieved 92% for accuracy and occult jaundice achieved 94% for accuracy.

Referring to Table 3, it can be found that accuracy distribution on each case appears substantially different in most model, except R-JaunLab. Xception with RAN yielded an accuracy for mean of 90.67%, 96%, 100% and 76% on photos of subjects in total average, healthy people, obvious jaundice and occult jaundice, respectively. ResNet50 with RAN yielded an average accuracy of 89.33%, 88%, 100% and 80% on testing images (for total average, healthy

Table 2 Results of multi-class recognition of comparative experiments on training subset, validation subset and testing subset of photos of subjects. 224 × 224 and 299 × 299 are default shape size of input resolution from relevant models

Network	Size		Training				Validation			
	224 × 224	299 × 229	Accuracy	Precision	Sensitivity	Specificity	Accuracy	Precision	Sensitivity	Specificity
RAN + Xception [71]		✓	0.8784±0.1291	0.9216	0.8098	0.9749	0.8329±0.0867	0.8705	0.7543	0.9506
RAN + ResNet50 [27]	✓		0.9061±0.1234	0.9150	0.8768	0.9715	0.8213±0.0683	0.8481	0.7920	0.9277
RAN + InceptionV3 [30]		✓	0.9390±0.0896	0.9556	0.9132	0.9829	0.8551±0.0517	0.8780	0.8290	0.9421
RAN + VGG19 [29]	✓		0.8876±0.1289	0.9034	0.8515	0.9660	0.7748±0.0728	0.8124	0.7527	0.9124
R-JaunLab	✓		0.9685±0.0639	0.9753	0.9573	0.9896	0.9006±0.0566	0.9103	0.8838	0.9575

Table 3 Summary of performance on testing subset for the healthy controls, obvious jaundice and occult jaundice. The recognition performance with optimal shape size of input resolution and total average accuracy is also summarized in the columns denoted as input size and total, respectively

Network	Input size	Testing accuracy			
		Total	Healthy	Obvious	Occult
RAN + Xception [71]	299 × 299	0.9067	0.96	1	0.76
RAN + ResNet50 [27]	224 × 224	0.8933	0.88	1	0.8
RAN + Inception V3 [30]	224 × 224	0.9133	0.86	1	0.88
RAN + VGG19 [29]	224 × 224	0.88	0.84	0.98	0.82
R-JaunLab	112 × 112	0.9533	0.92	1	0.94

'Healthy', 'Obvious' and 'Occult' represent three classes: healthy controls, obvious jaundice and occult jaundice, respectively

people, obvious jaundice and occult jaundice, respectively). InceptionV3 with RAN yielded on photos of subjects an accuracy for mean of 91.33% on total average, 86% on healthy people, 100% on obvious jaundice and 88% occult jaundice. VGG19 with RAN yielded on testing photos an average accuracy of 88%, 84%, 98% and 82% (for total average, healthy people, obvious jaundice and occult jaundice, respectively). Furthermore, these classes identified by experiment were also verified by conventional chemical and biological procedures of *TBIL*.

It is a great advance that the R-JaunLab annotates the eye region and recognizes jaundices on the photos of subjects. The RAN network of R-JaunLab preserves fully information of eye region on the photos of subjects and performs the patch extraction methods. Patch-based methods are popular occurrence [70, 80, 81] by using biomarker for eye region, which is region of interest (RoI) for jaundice and only a fraction of photos of subjects. However, it brings up an distinct obstacle, while non-jaundice patches will result in deviations of the parameter optimization and learning, i.e., the non-jaundice region biases the proposed model in the training phase. Therefore, only the proposal that cropped by the yellow boxes meets the requirements of our work. Hence, we carefully use photos of subjects as input of the R-Jaundice model, which is more accurate and improves the efficiency of clinical diagnosis as well as prognosis.

Multi-class recognition has more clinical values than binary-class recognition (i.e., healthy controls and obvious jaundice), because recognition of occult jaundice makes possible non-invasive diagnosis, which relieves the pressure of the patient and assists the skilled individuals to make more early optimal therapeutic procedure. In addition, CNNs have been sufficiently developed and also performed for biomedical image analysis, e.g., image classification [82, 83], image segmentation [84–86] and image registration [87–89], but there still exists a lot of improvement of biomedical image data in contrast to computer vision community [27, 90–93]. The proposed model employs an optimal training strategy (i.e., transfer learning based on VGG16), which fine-tunes the optimal parameters and converges quickly. This schema contributes more to our challenging task.

Conclusions

In summary, architecture of R-JaunLab holds the following benefits of 1) automatic extraction of eye region; 2) automatically enriched learning of simple information, obvious information and discriminative information through feature hierarchy (i.e., low-level feature, middle-level feature and high-level feature, respectively); 3) uncomplicated training process (end-to-end); 4) promising performance of fine-tuning; and 5) finally, computer-assisted diagnose of multi-class jaundice is non-invasive method, which is based on photos of subjects.

Jaundice is a common and complex clinical symptom with potential involvement in hepatology, general surgery, infectious diseases, pediatrics, genetic diseases, gynecology and obstetrics. Currently, jaundice is ascertained by the combination of inspection of doctors and laboratory test for serum bilirubin level. However, the accuracy of visual examination depends on the experience and subjective judgment of doctors. Although the test results of serum bilirubin are objective and reliable, patients can only be tested in the medical institutions where the test is carried out. Moreover, the diagnostic method based on blood test is inconvenient and time-consuming and increases the public health expenditure.

This study proposed a computer-aided diagnostic system of jaundice based on the sclera photos, which can give a reliable answer after the sclera photos are input. Thus, jaundice can be diagnosed intelligently by the system without relying on doctors and medical institutions. Intelligent diagnosis can not only make up for personal experience limitations and visual resolution limitations of doctors, but also bring convenience for patients far away from hospitals, which could increase diagnostic accuracy and efficacy. We are trying to make this diagnosis system into software that can be installed on smartphones. In the future, as long as you have a camera smartphone, you can complete the diagnosis.

Declarations

Conflicts of Interest The authors declare no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

1. Chambers CV, Irwin CE: Intense jaundice in an adolescent. an unusual presentation of infectious mononucleosis. J. Adolesc. Health Care, 7(3):195–197, 1986
2. Anand AC, Puri P: Jaundice in malaria. J Gastroenterol Hepatol, 20(9):1322–1332, 2010
3. Chan LY, Tsang WC, Hui Y, Leung WY, Chan KL, Sung JY: The role of lamivudine and predictors of mortality in severe flare-up of chronic hepatitis b with jaundice. J Viral Hepat, 9(6):424–428, 2010
4. Howard R, Watson CJ: Antecedent jaundice in cirrhosis of the liver. Arch Intern Med, 80(1):1–10, 1947
5. Hawkins WG, Dematteo RP, Jarnagin WR, Ben-Porat L, Fong Y: Jaundice predicts advanced disease and early mortality in patients with gallbladder cancer. Ann Surg Oncol, 11(3):310–315, 2004
6. Brandabur JJ, Kozarek RA, Ball TJ, Hofer BO, Jr RJ, Traverso LW, Freeny PC, Lewis GP: Nonoperative versus operative treatment of obstructive jaundice in pancreatic cancer: cost and survival analysis. Am J Gastroenterol, 83(10):1132, 1988
7. Larry C: Oxford textbook of primary medical care. J R Soc Med, 97(6):304, 2004
8. Maisels MJ: Managing the jaundiced newborn: a persistent challenge, CMAJ
9. Roche SP, Kobos R: Jaundice in the adult patient. Am Fam Physician, 69(2):299–304, 2004
10. Labori KJ, Raeder MG: Diagnostic approach to the patient with jaundice following trauma. Scandinavian Journal of Surgery Sjs Official Organ for the Finnish Surgical Society & the Scandinavian Surgical Society, 93(3):176, 2004
11. Winger J, Michelfelder A: Diagnostic approach to the patient with jaundice. Prim Care, 38(3):469–482, 2011
12. Aydım M, Hardala FC, Ural B, Karap S: Neonatal jaundice detection system, J Med Syst. 40(7):166, 2016
13. Halder A, Banerjee M, Singh S, Adhikari A, Sarkar PK, Bhattacharya AM, Chakrabarti P, Bhattacharyya D, Mallick AK, Pal SK: A novel whole spectrum-based non-invasive screening device for neonatal hyperbilirubinemia. IEEE J Biomed Health Inform, PP(99):1.
14. Mannino RG, Myers DR, Tyburski EA, Caruso C, Boudreaux J, Leong T, Clifford GD, Lam WA: Smartphone app for non-invasive detection of anemia using only patient-sourced photos. Nat Commun, 9(1), 2018
15. Padidar P, Shaker M, Amoozgar H, Khorraminejad-Shirazi M, Hemmati F, Najib KS, Pourarian S: Detection of neonatal jaundice by using an android os-based smartphone application. Iran J Pediatr, 29(2):e84397, 2019
16. Thompson BL, Wyckoff SL, Haverstick DM, Landers JP: Simple, reagentless quantification of total bilirubin in blood via microfluidic phototreatment and image analysis. Anal Chem, 89(5):3228–3234, 2017.
17. Saha S, Saha S, Bhattacharyya PP: Classifier fusion for liver function test based indian jaundice classification. In International Conference on Man & Machine Interfacing, 2016.
18. Saini N, Kumar A: Comparison of non-invasive bilirubin detection techniques for jaundice prediction, 2016.
19. Wang X, Zhang A, Han Y, Wang P, Sun H, Song G, Dong T, Yuan Y, Yuan X, Zhang M. Urine metabolomics analysis for biomarker discovery and detection of jaundice syndrome in patients with liver disease. Mol Cell Proteomics Mcp, 11(8):370, 2012
20. Zulkarnay Z, Jurimah AJ, Ibrahim B, Shazwani S, Nasir MAKA: An overview on jaundice measurement and application in biomedical: The potential of non-invasive method. In International Conference on Biomedical Engineering, 2015.
21. Knill-Jones RP, Stern RB, Girmes DH, Maxwell JD, Thompson RP, Williams R: Use of sequential bayesian model in diagnosis of jaundice by computer. Br Med J, 1(5852):530–533, 1973
22. Laddi A, Kumar S, Sharma S, Kumar A: Non-invasive jaundice detection using machine vision. IETE J Res, 59(5):591–596, 2013
23. Krizhevsky A, Sutskever I, Hinton G: Imagenet classification with deep convolutional neural networks. In International Conference on Neural Information Processing Systems, 2012
24. Kumar M, Dargon S: A survey of deep learning and its applications: A new paradigm to machine learning. Arch Comput Meth Eng, 2019
25. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD: Backpropagation applied to handwritten zip code recognition. Neural Comput, 1(4):541–551, 2014
26. Payal C, Kumar GN, Munish K. Content-based image retrieval system using orb and sift features. Neural Comput Applic, 32:2725–2733, 2020
27. He K, Zhang X, Ren S, Sun J: Deep residual learning for image recognition. In 2016 IEEE Conf Comp Vis Pattern Recognit (CVPR), pages 770–778, 2016
28. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ: Densely connected convolutional networks. In 2017 IEEE Conf Comput Visi Pattern Recognit (CVPR), pages 2261–2269, 2017
29. Simonyan K, Zisserman A: Very deep convolutional networks for large-scale image recognition. Computer Science, 2014
30. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z: Rethinking the inception architecture for computer vision. 2015
31. Girshick R: Fast r-cnn. In 2015 IEEE International Conference on Computer Vision (ICCV), pages 1440–1448, 2015
32. Girshick R, Donahue J, Darrelland T, Malik J: Rich feature hierarchies for object detection and semantic segmentation. In IEEE Conference on Computer Vision & Pattern Recognition, 2014
33. He K, Georgia G, Piotr D, Ross G: Mask r-cnn. IEEE Trans Pattern Anal Mach Intell, PP(99):1, 2017
34. He K, Zhang X, Ren S, Sun J: Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans Pattern Anal Mach Intell, 37(9):1904–16, 2014
35. Jung C, Sun T, Jiao L: Eye detection under varying illumination using the retinex theory. Neurocomputing, 113(596):130–137, 2013
36. Ren S, He K, Girshick R, Sun J. Faster r-cnn: Towards real-time object detection with region proposal networks. IEEE Trans Patt Anal Mach Intell, 39(6):1137–1149, 2017
37. Soetedjo A: Eye detection based-on color and shape features. Int J Adv Comput Sci Appl, 3(5), 2012
38. Kumar M, Bansal M, Kumar M. 2d object recognition techniques: State-of-the-art work. Archives of Computational Methods in Engineering, 02 2020
39. Kumar M, Dargan S: A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. Expert Systems with Applications, pages 1–27, 11 2019

40. Kumar M, Gupta S, Thakur K: 2d-human face recognition using sift and surf descriptors of face's feature regions. Vis Comput, 01 2020

41. Kumar M, Kumar R, Kaur P: A healthcare monitoring system using random forest and internet of things (iot). Multimed Tools Appl, 02 2019

42. Bland JM, Altman D: Measuring agreement in method comparison studies. Stat Methods Med Res, 8:135–60, 1999

43. PS Myles, Cui JI: using the bland altman method to measure agreement with repeated measures. Br J Anaesth, 99(3):309–311, 2007

44. Ronneberger O, Fischer P, Brox T: U-net: Convolutional networks for biomedical image segmentation. 2015

45. Nair V, Hinton GE: Rectified Linear Units Improve Restricted Boltzmann Machines. In International Conference on International Conference on Machine Learning, 2010

46. Goroshin R, Mathieu M, LeCun Y: Learning to Linearize Under Uncertainty. CoRR, abs/1506.03011, 2015

47. Creswell A, Arulkumaran K, Bharath AA: On denoising autoencoders trained to minimise binary cross-entropy. 2017

48. Deng J, Dong W, Socher R, Li LJ, Li FF: Imagenet: a large-scale hierarchical image database. In IEEE Conference on Computer Vision & Pattern Recognition, 2009

49. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M: Imagenet large scale visual recognition challenge. Int J Comput Vis, 115(3):211–252, 2015

50. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S: Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans Med Imaging, 35(5):1207–1216, 2016

51. Cho J, Lee K, Shin E, Choy G, Do S: How much data is needed to train a medical image deep learning system to achieve necessary high accuracy? Computer Science, 2015

52. Greenspan H, van Ginneken B, Summers RM: Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. IEEE Trans Med Imaging, 35(5):1153–1159, 2016

53. Rajkomar A, Lingam S, Taylor AG, Blum M, Mongan J: High-throughput classification of radiographs using deep convolutional neural networks. J Dig Imaging, 30(1):95–101, 2017

54. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM: Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging, 35(5):1285–1298, 2016

55. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J. Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Trans Med Imaging, 35(5):1299–1312, 2016

56. Wang Z, Meng Y, Weng F, Chen Y, Lu F, Liu X, Hou M, Zhang J: An effective cnn method for fully automated segmenting subcutaneous and visceral adipose tissue on ct scans. Ann Biomed Eng, pages 1–17, 2019

57. KA: Cs231n course notes: Transfer learning [online]. Accessed: 19-May-2016 http://cs231n.github.io/transfer-learning.

58. Razavian AS, Azizpour H, Sullivan J, Carlsson S: Cnn features off-the-shelf: An astounding baseline for recognition. pages 512–519, 2014

59. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A: Learning deep features for discriminative localization. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2921–2929, 2016

60. Zeiler MD, Fergus R: Visualizing understanding convolutional networks. 2013

61. Dosovitskiy A, Springenberg JT, Riedmiller M, Brox T: Discriminative unsupervised feature learning with convolutional neural networks. 2014

62. Wong SC, Gatt A, Stamatescu V, Mcdonnell MD: Understanding data augmentation for classification: When to warp? 2016

63. Hinton G, Tieleman T: Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. COURSERA: Neural Networks for Machine Learning, (4):26–30, 2012

64. Bengio Y: Speeding up stochastic gradient descent. 2007

65. Kaleem, Rashid, Sreepathi, Pingali, Keshav. Stochastic gradient descent on gpus. 2015

66. Dutta P, Saha S, Gulati S: Graph-based hub gene selection technique using protein interaction information: Application to sample classification. IEEE J Biomed Health Inform, PP(99):1, 2019

67. Ji J, Zhang A, Liu C, Quan X, Liu Z: Survey: Functional module detection from protein-protein interaction networks. IEEE Trans Knowl Data Eng, 26(2):261–277, 2013

68. Jr MA, Niswender GD, Rebar RW. Principles for the assessment of the reliability of radioimmunoassay methods (precision, accuracy, sensitivity, specificity). Acta Endocrinologica Supplementum, 142(1 Suppl):163, 1969

69. Kessler RC, Abelson JM, Demler O, Escobar JI, Zheng H: Clinical calibration of dsm-iv diagnoses in the world mental health (wmh) version of the world health organization (who) composite international diagnostic interview (cidi). 13(2):122–139, 2004

70. Spanhol FA, Oliveira LS, Petitjean C, Heutte L: Breast cancer histopathological image classification using convolutional neural networks. In International Joint Conference on Neural Networks, 2016

71. Chollet F: Xception: Deep learning with depthwise separable convolutions. In 2017 IEEE Conf Comput Vis Pattern Recognit (CVPR), pages 1800–1807, 2017

72. Kouvaris K, Clune J, Kounios L, Brede M, Watson RA: How evolution learns to generalise: Principles of under-fitting, over-fitting and induction in the evolution of developmental organisation. Journal of the Society of English & American Literature Kansei Gakuin University, 52(4):93–107, 2015

73. Cordero C, Schieve LA, Croen LA, Engel SM, Maria ASR, Herring AH, Vladutiu CJ, Seashore CJ, Daniels JL: Neonatal jaundice in association with autism spectrum disorder and developmental disorder. Journal of perinatology: official journal of the California Perinatal Association, 2019

74. Redfern V, Mortimore G: Right hypochondrial pain leading to diagnosis of cholestatic jaundice and cholecystitis: a review and case study. Gastrointestinal Nursing

75. Xu X, Zhang X: The application of intravoxel incoherent motion diffusion-weighted imaging in the diagnosis of hilar obstructive jaundice. J Comput Assist Tomogr, 43(2):1, 2019

76. Tabatabaee RS, Golmohammadi H, Ahmadi SH: Easy diagnosis of jaundice: A smartphone-based nanosensor bioplatform using photoluminescent bacterial nanopaper for point-of-care diagnosis of hyperbilirubinemia. ACS sensors, 4(4):1063–1071, 2019

77. Tibana TK, Grubert RM, Fornazari VAV, Barbosa FCP, Bacelar B, Oliveira AF, Marchiori E, Nunes TF: The role of percutaneous transhepatic biliary biopsy in the diagnosis of patients with obstructive jaundice: an initial experience. Radiologia Brasileira, (AHEAD), 2019

78. Sunwoo MH, Lee JW, Kim JH: Method and apparatus for jaundice diagnosis based on an image, Apr. 18 2019. US Patent App. 16/115,821

79. Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network. Computer Science, 14(7):38–39, 2015

80. Litjens G, Sanchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, Hulsbergenvan DKC, Bult P, Van GB, Van DLJ: Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci Rep, 6(1):26286, 2016

81. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH: Deep learning for identifying metastatic breast cancer. 2016

82. Han Z, Wei B, Zheng Y, Yin Y, Li K, Li S. Breast cancer multi-classification from histopathological images with structured deep learning model. Scientific Reports, 7(1):4172, 2017

83. Takiyama H, Ozawa T, Ishihara S, Fujishiro M, Shichijo S, Nomura S, Miura M, Tada T: Automatic anatomical classification of esophagogastroduodenoscopy images using deep convolutional neural networks. Sci Rep, 8(7497):7497, 2018

84. Causey JL, Zhang J, Ma S, Jiang B, Qualls JA, Politte DG, Prior F, Zhang S, Huang X: Highly accurate model for prediction of lung nodule malignancy with ct scans. Sci Rep, 8(1):9286, 2018

85. Deniz CM, Hallyburton S, Welbeck A, Honig S, Cho K, Chang G: Segmentation of the proximal femur from mr images using deep convolutional neural networks. Sci Rep, 8(1), 2018

86. Ghafoorian M, Karssemeijer N, Heskes T, Uden IWM, Sanchez CI, Litjens G, Leeuw FE, Ginneken B, Marchiori E, Platel B: Location sensitive deep convolutional neural networks for segmentation of white matter hyperintensities. Sci Rep, 7(1):5110, 2017

87. Hao C, Qi D, Xi W, Jing Q, Heng PA: Mitosis detection in breast cancer histology images via deep cascaded networks. In Thirtieth Aaai Conference on Artificial Intelligence, 2016

88. Suk HI, Lee SW, Shen D: Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis. Neuroimage, 101:569–582, 2014

89. Wang S, Kim M, Wu G, Shen D: Chapter11c scalable high performance image registration framework by unsupervised deep feature representations learning. IEEE Trans Biomed Eng, 63(7):1505–1516, 2016

90. Hinton GE, Salakhutdinov RR: Reducing the dimensionality of data with neural networks. Science, 313(5786):504–507, 2006

91. Lecun Y, Bengio Y, Hinton G: Deep learning. Nature, 521(7553):436, 2015

92. Shen D, Wu G, Suk HI: Deep learning in medical image analysis. Annu Rev Biomed Eng, 19(1):221–248, 2017

93. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A: Going deeper with convolutions. In 2015 IEEE Conf Comput Vis Pattern Recognit (CVPR), pages 1–9, 2015

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.